pennsylvania
DEPARTMENT OF EDUCATION

Response to PVAAS Misconceptions: District/School Reporting

Have you heard conflicting information on what PVAAS is, how it's used, and whether or not it provides reliable information?

Are there myths and misconceptions about PVAAS that raise questions for you?

This document has been created to identify important issues often raised regarding PVAAS reporting and to provide clarification about these misconceptions.

Achievement and Growth

PVAAS provides an objective and reliable way to measure the academic growth of groups of students and the value teachers, schools and districts add to students' educational experiences. PVAAS is one of the data tools in the cadre of tools provided to Local Education Agencies (LEAs) from the Pennsylvania Department of Education. Districts and schools are using PVAAS, in conjunction with achievement data, to evaluate the growth made by groups of students and to make sure all students are on the trajectory to proficiency. Utilizing all the data available (both achievement and growth), educators are able to make better data-informed instructional decisions to ensure the academic growth and achievement of all students.

PVAAS provides feedback to key questions:

1. Did each grade-level or course-level student group meet the standard for PA Academic Growth in English/Language Arts (grades 4-8), Math (grades 4-8), Science (grades 4 and 8), Algebra I, Literature, and Biology?
2. Did subgroups of students at each grade level or in each Keystone content area meet the standard for PA Academic Growth in English/Language Arts (grades 4-8), Math (grades 4-8), Science (grades 4 and 8), Algebra I, Literature, and Biology
3. Is each individual student on a trajectory to success on future assessments, such as PSSA, Keystone, PSAT, ACT and SAT assessments?

Achievement results (PSSA and Keystone) and growth results (PVAAS) must be used together to get a complete picture of student learning. To view the achievement results of Pennsylvania's public districts/schools, please visit the PDE website at http://www.pde.state.pa.us. To view the growth results of Pennsylvania's public districts/schools, please go to: https://pvaas.sas.com.

Disclaimer for ANY Data Tool

- NO data source should ever be considered in isolation.
- ALL educational decisions should be made on the basis of multiples sources of data.
- ALL data provide *indicators* of phenomena.

When new data are gathered, the educational professional should ask:

- Do these data provide insights that have not been available before?
- Are these data consistent with data already collected?
- Do these data confirm or conflict with our existing profile of our students or our educational programs?
- What other data should be investigated in light of the new insights?
- What questions do we have about our standards-aligned system based on these data?
- What questions do we have about student performance based on these data?

PVAAS Misconceptions: These statements are NOT accurate about PVAAS.

1. Growth is always highly correlated with certain student demographic variables. Hence, PVAAS should provide direct, additional controls/adjustments for demographics.
2. PVAAS is not reliable or valid since it is only based on the PSSA.
3. PVAAS is based on a 'black box' methodology.
4. The PVAAS methodology is too complex; a more simple approach to measuring district and school effectiveness would provide better information to educators.
5. The PSSA and Keystone exams are not designed to discriminate well at the extremes so growth cannot be calculated using these assessments.
6. If students are already high achieving, it is harder to show growth.
7. It is not possible to show growth with all groups of students, such as students with IEPs or students identified as gifted.
8. PVAAS should always indicate growth if our percent of students proficient/advanced increased since last year.
9. PVAAS cannot measure the academic growth of districts and schools with high mobility rates.
10. PVAAS cannot measure growth for groups of students who have missing data.

Misconception #1: Growth is always highly correlated with certain demographic variables. Hence, PVAAS should provide direct, additional controls/adjustments for demographics.

There is typically little or no relationship between growth and demographic variables, such as socioeconomic status.

From the 1960s, it has been established that there is a relationship between student demographics and achievement results. This is often not the case when considering growth! A review of the literature indicates there is typically little or no relationship between the academic growth of students and those students' demographic factors, such as socioeconomic status and racial/ethnic background, as long as the prior achievement of those students is sufficiently accounted for. Some key research findings include:

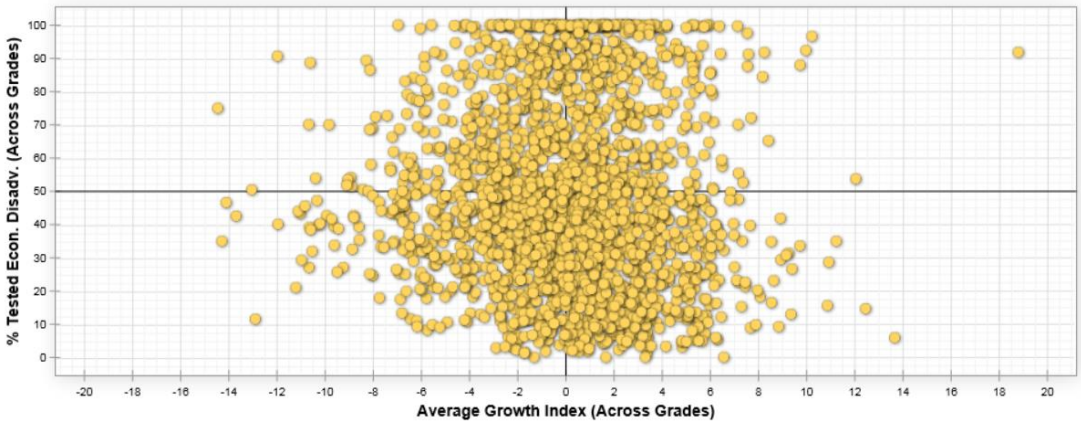
- A single measure of student achievement has inherent limitations due to the fact that achievement is correlated to a student's socioeconomic status and past performance (Hershberg, et al.; Olson, 2007; Sanders, 2000).
- Fallon (2004) reports that the importance of value-added assessment is based on the experimental design that removes virtually all influence of genetics and socio-economic factors. The design provides a measure of the direct effect of classroom instruction and the instructional program.
- [Value-added assessment systems] can remove the effects of factors not under the control of the school, such as prior performance and socioeconomic status, and thereby provides a more accurate indicator of school or teacher influence than is possible when these factors are not controlled (McCaffrey, Lockwood, Koretz & Hamilton, 2003; Ross, Wang, Sanders, Wright & Stringfield, 1999a; Wright, Horn & Sanders, 1997).

PVAAS does not control for demographics as its methodology uses all prior historical data to sufficiently account for measurement error. PVAAS compares a group of students in a school to themselves; hence the demographics have remained consistent and are represented in the longitudinal assessment history of each student. The students were more than likely representative of that "subgroup" when they took their prior assessments. Given that PVAAS uses all longitudinal data to yield measures of growth, the subgroup variable is already represented in the analysis. Their membership in a subgroup existed when they took their prior assessments.

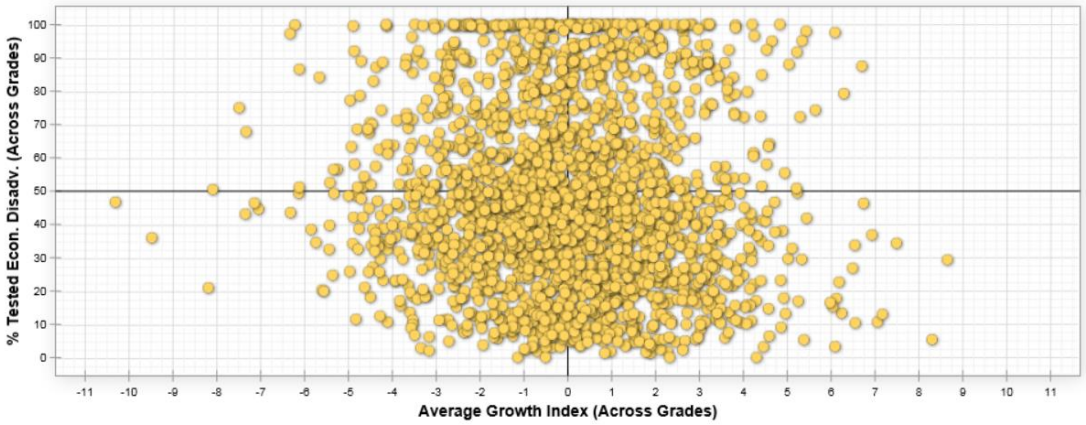
Evidence of the research findings exist in actual PA data. Evidence from PVAAS reporting in Pennsylvania has yielded results to show that there are schools in Pennsylvania making significant academic growth with students with IEPs, ELL students, minority students, and economically disadvantaged students. In fact most districts/schools in Pennsylvania can find this evidence in their own PVAAS reporting. The statewide PVAAS scatterplots (available on PDE's PVAAS webpage) illustrate evidence to show there are many high achieving schools making high growth AND many low achieving schools making high growth in most of the tested subjects and grades.

The statewide scatter plots displaying PVAAS growth indices versus the percentages of students served who are economically disadvantaged are provided here.

MATH



READING



There are a few subjects and grades (namely, PSSA Science and the Keystones) where there is a small or moderate relationship between growth and students from certain subgroups. In interpreting these results, it must be emphasized that, *at the individual student-level and taking into account a student's prior testing history, characteristics like the socioeconomic status of that student does not have a relationship to the student's ability to show growth*. Data from other states indicates that there is typically no relationship between growth measures and demographics, even in the end-of-course or science assessments. With that in mind, the results of this reporting in PSSA Science and the Keystones is an opportunity to re-assess whether the standards-aligned system is fully implemented in all classrooms, schools and districts and whether there are additional needs and supports for certain student populations, schools and districts. PDE will continue to monitor results.

References

Olson, A. (2007). Growth measures for systemic change. *School Administrator*. *64*(1), 10-15.

Fallon, D. (2004). *Clarifying how we think about teaching and learning*. Proceeding from the 2004 National Value-Added Conference. Columbus, OH: Battelle for Kids.

Hershberg, T., Simon, V. A., & Lea-Kruger, B. (2004) Measuring what matters: How value-added assessment can be used to drive learning gains. *American School Board Journal, 2,* 27-31.

McCaffrey, D., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Models for value-added modeling of teacher effects.* Santa Monica, CA: RAND Corporation.

Ross, S. M., Sanders, W. L., & Wright S.P. (2000a). *Fourth-year achievement results on the Tennessee Value-Added Assessment System for restructuring schools in Memphis.* TN: University of Memphis.

Sanders, W. L. (2000). Value-added assessment from student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education*, *14*(4), 329-339.

Wright, P. S., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel in Education, 11*, 57-67.

Misconception #2: PVAAS is not reliable or valid since it is only based on state assessments.

The PSSA forms the foundation of longitudinal assessment data for PVAAS analyses in Mathematics, English/Language Arts, Science and the Keystones. The quality of the PSSA is assessed and ensured in various ways.

Pennsylvania is continuously required to ensure the quality of the PSSA. In 2004 an independent evaluation of the PSSA was conducted to respond to questions such as:

1. Does the PSSA adequately measure the academic content specified by the State Standards contained in Chapter 4?
2. Are the PSSA tests internally consistent and replicable?
3. Do the scores produced by the PSSA correlate positively and significantly with pertinent scores produced on related tests, such as Terra Nova, Stanford Achievement Test, etc.?
4. Were the methodologies used to determine performance levels (cut scores) reasonable and technically competent?

This study was conducted by the Human Resources Research Organization (HumRRO). These results can be found on PDE's website.

Additionally, Pennsylvania's PSSA receives rigorous review by the federal government, as well as a statewide advisory group of experts on national assessments, psychometrics, and statistics to ensure a quality assessment in Pennsylvania. The rigorous psychometrics on the PSSA can be found in the annually released PSSA Technical Report on PDE's website (http://www.pde.state.pa.us).

PVAAS GROWTH MEASURES

- The PVAAS growth methodologies are based on more than 30 years of research and experience.
- The PVAAS methodologies and algorithms have been discussed and published in a book titled, "Grading Teachers, Grading Schools" edited by Jason Millman.
    - Chapters 12 through 16 focus on the Education Value-Added Assessment System (EVAAS) upon which PVAAS is based.
- PVAAS uses a robust, multivariate, longitudinal mixed effect model in its analyses to yield quality measures of growth. It is NOT a simple comparison of two test scores!
    - ALL available prior PSSA assessment scores are used in the analyses.
    - Standard errors are always reported and considered in the analyses and reporting.
- The following research from the RAND Corporation corroborates the SAS EVAAS for K-12 modeling approaches.
    - On the **choice of a complex value-added model**: McCaffrey, D. F., Han, B. and Lockwood, J. R. (2008). "Value-Added Models: Analytic Issues." A paper prepared for the National Research Council and the National Academy of Education, Board on Testing and Accountability Workshop on Value-Added Modeling, Nov. 13-14, 2008, Washington D.C.
    - On the **advantages of the longitudinal, mixed model approach**: Lockwood J.R. and McCaffrey D.F. (2007). "Controlling for Individual Heterogeneity in Longitudinal Models, with Applications to Student Achievement." Electronic Journal of Statistics, Vol. 1, 223-252.
    - On the **insufficiency of simple value-added models**: McCaffrey, D. F., Han, B. and Lockwood, J. R. (2008). "From Data to Bonuses: A Case Study of the Issues Related to Awarding Teachers Pay on the Basis of the Students' Progress." A paper presented at the conference on Performance Incentives: Their Growing Impact on American K-12 Education, Feb. 28-29, 2008, National Center on Performance Incentives at Vanderbilt University.
- The National Governors Association included the SAS EVAAS for K-12 application in its 2003 Data-Driven Decision Making Tool Kit.

Misconception #2 (continued):

PVAAS PROJECTION DATA

- PVAAS student-level projection data are reliable, and are created using a model that has been reviewed and approved by four different review panels and the GAO (US Government Accountability Office).
    - The 2008 growth model proposal to USDOE includes information regarding the statistical model and projection reliability study.
        - http://www.ed.gov/admins/lead/account/growthmodel/pa/index.html
    - Research studies have confirmed that the PVAAS projections are more reliable at looking at the future performance of students than their most recent PSSA score.
- The PSSA and Keystones meets the three conditions to be used in PVAAS analyses.
    - Must be aligned to curriculum standards.
    - Must be reliable and valid.
    - Must demonstrate sufficient stretch at the extremes.
- The PSSA and Keystones are statewide assessments in Pennsylvania that perform a universal assessment of PA standards.
    - The PSSA exams are aligned to the appropriate grade level standards that are sufficient for longitudinal modeling and prediction.
- The SAS EVAAS team performs routine checks every year to look at the stretch and stability of the scales. To look at stretch, they do two things. First, they ensure there is a sufficient number of different scale scores at the top and bottom of the scales to differentiate student achievement. The SAS EVAAS team then looks at the percentage of students scoring at the top to ensure there are no ceilings.
- Stability of scales is also monitored by looking at the state distributions of scale scores every year to determine if the reliability and validity requirement is met; this has been satisfied each year.

Since these conditions have been met, the power of using PSSA and Keystone data lies in the fact that there are many students who take each of the PSSA exams each year. This has resulted in a very robust database of longitudinal students' performance results.

Misconception #3: PVAAS is based on a 'black box' methodology.

The EVAAS methodologies and algorithms are published and available for those interested in learning more about the statistical models used in all applications including Pennsylvania.

- On the **Tennessee Value-Added Assessment System:** Millman, J. (ed.) (1997). "Grading Teachers, Grading Schools: Is Student Achievement a Valid Evaluation Measure?"
    - Chapters 12 through 16 focus on the Education Value-Added Assessment System (EVAAS) upon which PVAAS is based.
- On the **SAS EVAAS Statistical Models upon which PVAAS is based**: SAS Institute Inc. (2015) "SAS® EVAAS® for K-12 Statistical Models."

There are multiple statistical models used in EVAAS analyses based on the objectives of the analyses and the characteristics and availability of the assessment data used.

- The multivariate response model (MRM) used in value-added analyses is a robust, multivariate, longitudinal mixed model. In other words, it is conceptually a multivariate repeated-measures ANOVA model.
    - The MRM is used when scores are scaled or transformed so that the difference between two scores is meaningful.
    - The MRM is used when there are clear 'before' and 'after' assessments in which to form a reliable growth estimate. In Pennsylvania, this is used for Math and Reading in grades 4-8.
- The univariate response model (URM) used in value-added analyses is conceptually an analysis of covariance (ANCOVA) model.
    - The URM is used when the test data do not meet the requirements for MRM analyses as stated above. In Pennsylvania, this is used in subjects where the PSSA is not tested in consecutive grades (Science, grades 4 and 8; or where students are tested in Keystone content area with the corresponding Keystone exam (Algebra I, Literature, and Biology).
- It is NOT a simple comparison of two test scores!
    - ALL available prior PSSA assessment scores from ALL students are used in the analyses, not simply the last two years!
    - Standard errors are always reported and considered in the analyses and reporting.

Misconception #4: The PVAAS methodology is too complex; a more simple approach to measuring district and school effectiveness would provide better information to educators.

PVAAS analyses and reporting are designed to provide measures of district and school effectiveness!

- Douglas Harris, in his recent 2011 book 'Value-Added Measures in Education: What Every Educator Needs to Know,' states that "basic value-added is really just simple growth and growth models are usually not considered to be value-added measures because they do not take into account as much information as they could." He goes on to state that "*describing* individual student growth is quite different (and easier) from the challenging task of measuring, with value-added approaches, how much individual teachers and schools contributed to that growth."

PVAAS value-added measures are NOT simple comparisons of test scores! Rather, PVAAS is based upon a univariate response model (URM/conceptually an analysis of covariance (ANCOVA) model) and a multivariate response model (MRM), which is a very robust, multivariate, longitudinal mixed statistical model. With these models, the entire history of students' test scores is used, not simply the last two years!

The following research from the RAND Corporation corroborates the SAS EVAAS for K-12 modeling approaches.

- On the **choice of a complex value-added model**: McCaffrey, D. F., Han, B. and Lockwood, J. R. (2008). "Value-Added Models: Analytic Issues." A paper prepared for the National Research Council and the National Academy of Education, Board on Testing and Accountability Workshop on Value-Added Modeling, Nov. 13-14, 2008, Washington D.C.
- On the **advantages of the longitudinal, mixed model approach**: Lockwood J.R. and McCaffrey D.F. (2007). "Controlling for Individual Heterogeneity in Longitudinal Models, with Applications to Student Achievement." Electronic Journal of Statistics, Vol. 1, 223-252.
- On the **insufficiency of simple value-added models**: McCaffrey, D. F., Han, B. and Lockwood, J. R. (2008). "From Data to Bonuses: A Case Study of the Issues Related to Awarding Teachers Pay on the Basis of the Students' Progress." A paper presented at the conference on Performance Incentives: Their Growing Impact on American K-12 Education, Feb. 28-29, 2008, National Center on Performance Incentives at Vanderbilt University.

Significant professional development has been conducted in Pennsylvania to create a conceptual understanding of the PVAAS methodology and the concept of growth as measured by PVAAS – how the value-added results are yielded. Given that a small proportion of people have training in longitudinal statistical modeling (not statistics), it is necessary to use conceptual approaches in professional development and resource materials.

Misconception #5: The PSSA and Keystone exams are not designed to discriminate well at the extremes so growth cannot be calculated using these assessments.

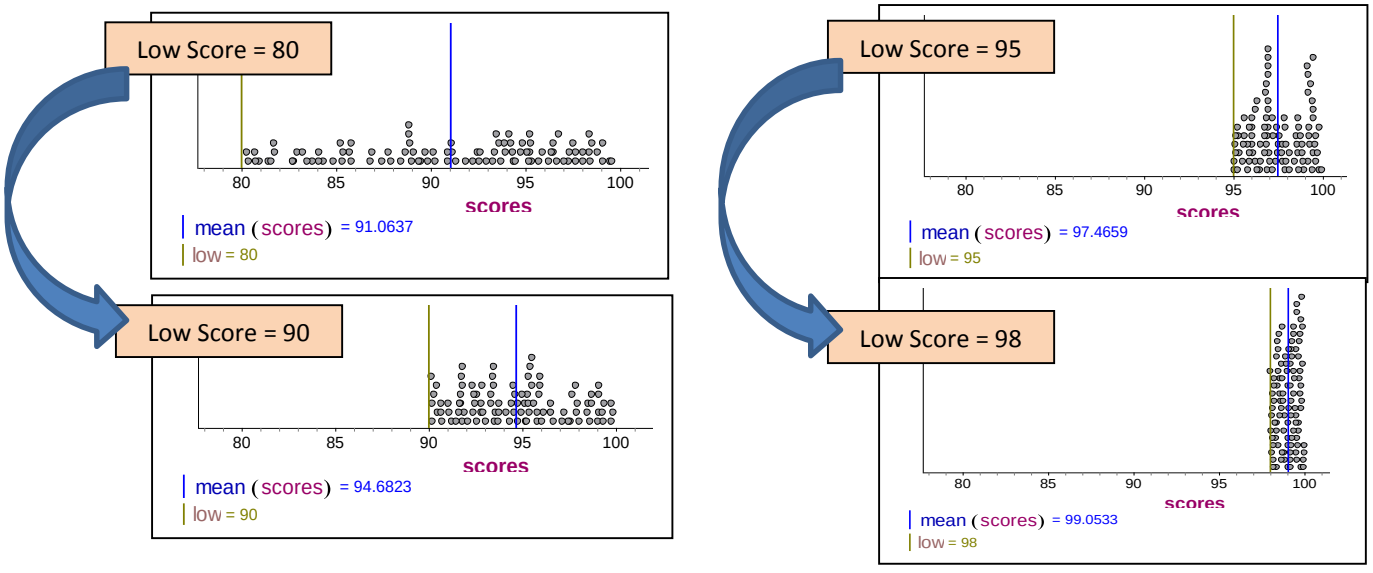
While Pennsylvania's state assessments are designed to discriminate proficient from non-proficient, they are also designed to have sufficient stretch to discriminate between Below Basic, Basic, Proficient, and Advanced performance levels!

- There is no ceiling on the PSSA or Keystones!
  - This can be verified through the Bureau of Assessment and Accountability, PA Department of Education.
  - Each year, test scores are scaled using the lowest score of 700 and the previous year's cut scores for the proficient category.
  - The high end is then allowed to be scaled based on the distribution of the data – not on a fixed, pre-determined value.
- Pennsylvania's state assessments meet the three conditions each year to be used in PVAAS analyses.
  - Must be aligned to curriculum standards.
  - Must be reliable and valid.
  - Must demonstrate sufficient stretch at the extremes.
- The SAS EVAAS team performs routine checks every year to look at the stretch and stability of the scales. To look at stretch, they do two things. First, they ensure there are enough different scale scores at the top and bottom of the scales to differentiate student achievement. The SAS EVAAS team then looks at the percentage of students scoring at the top to ensure there are no ceilings. After the analysis is done, they ensure that schools that serve both high and low achieving students can show both high and low growth. This is always verified each year!

High performing schools can and are showing growth in Pennsylvania! For example, when calculating growth for groups of students in Math and English/Language Arts, a value-added growth measure is the difference of the estimated mean (or average) performances from two consecutive school years. Watch how improvement of performance affects mean performance.

Example and Simulation – Math and ELA Example

- Suppose we have a population of high performing students whose mean (average) performances range from 80 to 100 (100 being the highest value).
- What happens to the mean (blue vertical bar) of these performances as we endeavor to move all students to higher performance levels? Notice that the mean (average) for each group increases as students move to higher performance levels; however, there is always room for the mean to go up or increase even higher!



Misconception #6: If students are already high achieving, it is harder to show growth.

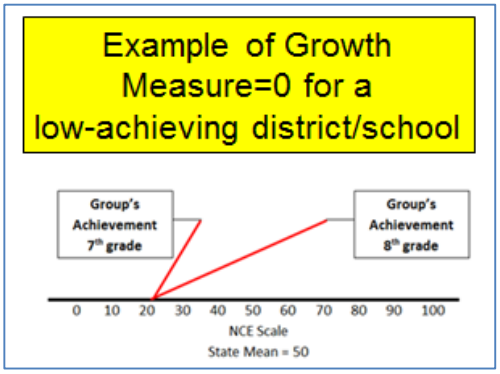
All students can meet the standard for PA Academic Growth as measured by value-added analyses!

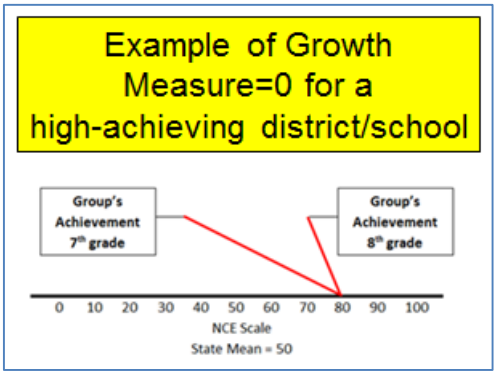
In PVAAS, the standard for PA Academic Growth is about maintaining achievement levels (English/Language Arts and Math, grades 4-8) or meeting expected performance (Science and Keystone content areas) based on a specific group's prior academic performance!

- If students do not meet the standard for PA Academic Growth, achievement results may be impacted.
- For low-achieving schools, simply meeting the standard for PA Academic Growth may not be sufficient or acceptable in order for students to meet long-term achievement goals of proficiency.
- For high-achieving groups, meeting (not exceeding) the standard for PA Academic Growth may be sufficient or acceptable.

PVAAS (English/Language Arts and Math, grades 4-8)

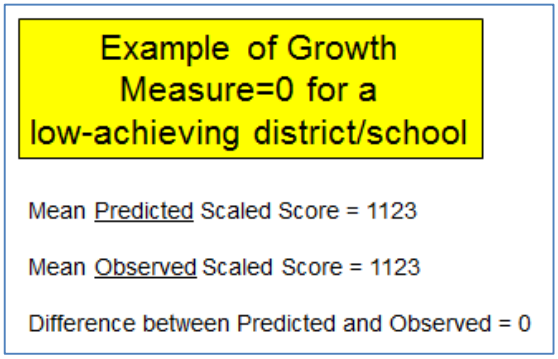
- A group of students meets the standard for PA Academic Growth (or, makes one year's worth of academic growth) when the student group maintains their relative achievement level from one year to the next.
- Growth measure estimates are based on longitudinal data, not single scores.
- In other words, a group of students makes one year's worth of academic growth in a school year when the group of students maintains its achievement position relative to the statewide distribution of scores each year.

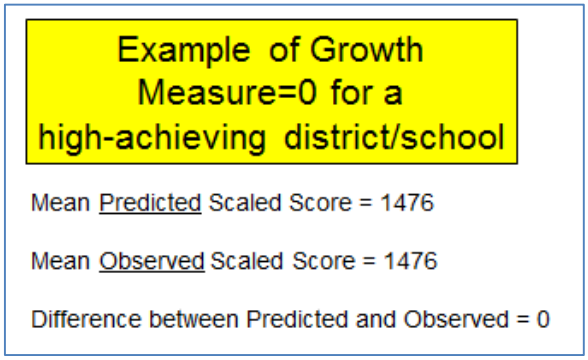




PVAAS (Science and Keystone content areas)

- A group of students meets the standard for PA Academic Growth (or, makes one year's worth of academic growth) when the mean observed score from the actual test is NOT significantly different from the mean predicted score for the student group.
- Predicted performance is based on longitudinal data, NOT single scores.
- In other words, a group of students makes one year's worth of academic growth in a school year when the students' actual performance is as expected.



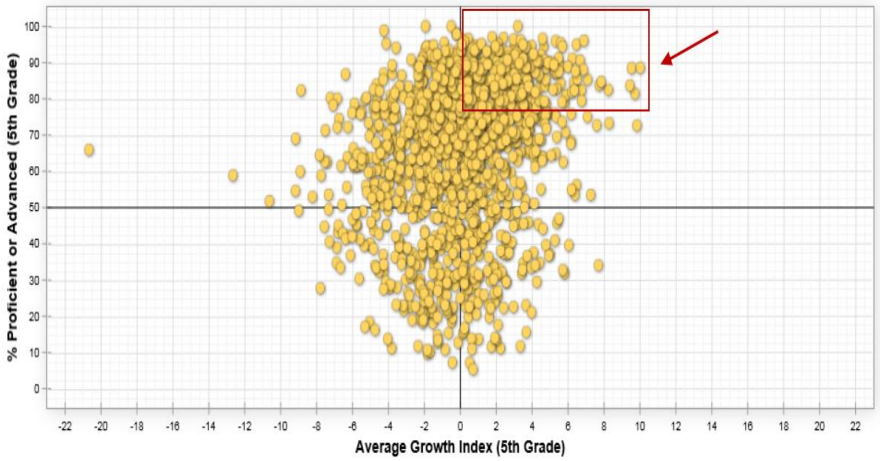


Misconception #6 (continued):

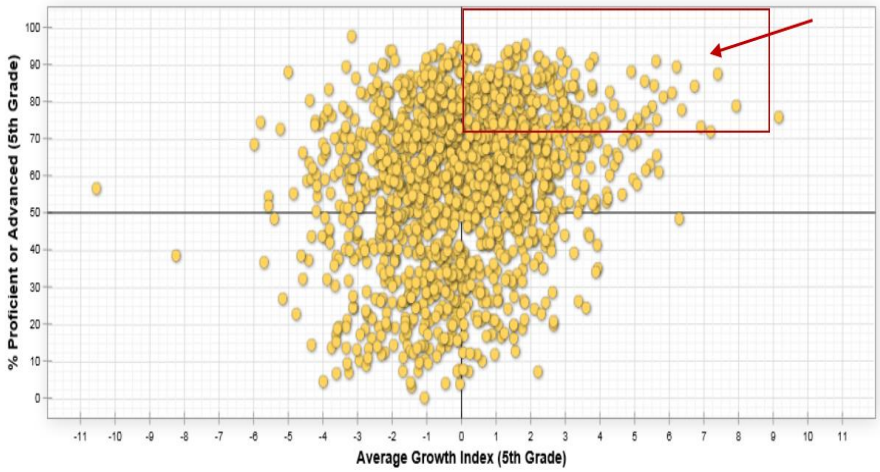
The examination of scatter plots (global change) provides further evidence that high achievement and high growth are NOT inconsistent. Statewide scatter plots in Pennsylvania for each grade level reveal that both high-achieving schools and low-achieving schools can and are making high growth at all grade levels!

Consider these two examples: 2014 5th Grade Math and 2014 5th Grade Reading. In both cases, we see there are many high performing schools that are exhibiting high growth as well. Refer to the red arrows!

MATH



READING



Pennsylvania has years of evidence to show that there are high achieving schools in Pennsylvania making high growth; there are low achieving schools in Pennsylvania making high growth; there are high achieving schools in Pennsylvania making low growth; and there are low achieving schools in Pennsylvania making low growth. We have evidence across the state of these scenarios in each grade and subject tested. While achievement results may be highly correlated to demographics, growth results are typically not!

What is most important to understand is that it is critical to have both types of data! Achievement data – tells you where the students are at a point in time AND Growth data –tells you how far the students have grown. Achievement and Growth - together they provide the more complete picture of student performance.

Misconception #7: It is not possible to show growth with all groups of students, such as students with IEPs or students identified as gifted.

While Pennsylvania's state assessments are designed to discriminate proficient from non-proficient, they are also designed to have sufficient stretch to discriminate between Below Basic, Basic, Proficient, and Advanced performance levels!

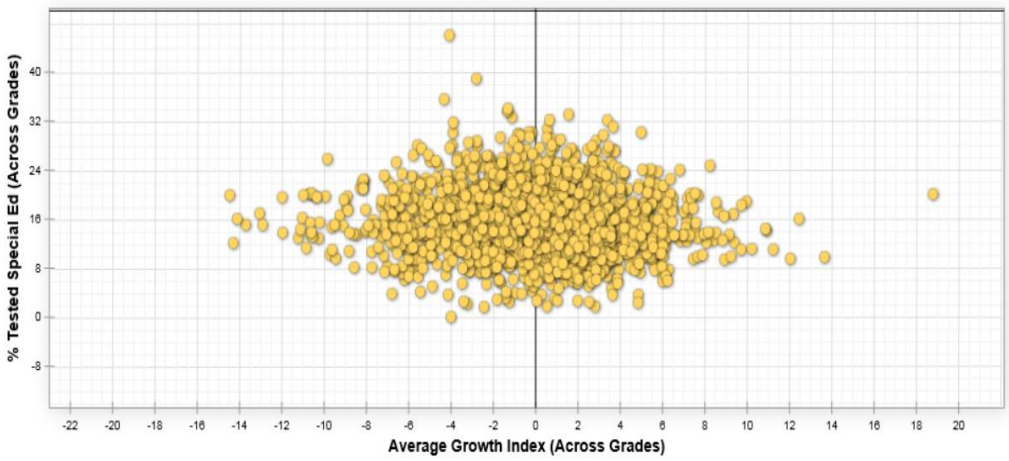
- Pennsylvania's state assessments meet the three conditions to be used in PVAAS analyses.
  - Must be aligned to curriculum standards.
  - Must be reliable and valid.
  - Must demonstrate sufficient stretch at the extremes.
- The SAS EVAAS team performs routine checks every year to look at the stretch and stability of the scales. To look at stretch, they do two things. First, they ensure there are enough different scale scores at the top and bottom of the scales to differentiate student achievement. They then look at the percentage of students scoring at the top to ensure there are no ceilings. After the analysis is done, the SAS EVAAS team ensures that schools that serve both high and low achieving students can show both high and low growth. This is always verified each year!
- If assessments have enough "stretch" to measure the achievement of both low- and high-achieving students, it is possible to measure the growth of all groups of students. Pennsylvania's state assessments meet the criteria!
- It is also important to note that students with IEPs and students identified as gifted may present a range of performance levels. One should not assume that all students with IEPS have low achievement results. As well, one should not assume that all students identified as Gifted are all high achieving.

PVAAS measures growth from the end of one year to the end of the next year, regardless of whether a student performs below, at, or above grade level…or performs at a Below Basic, Basic, Proficient or Advanced level.

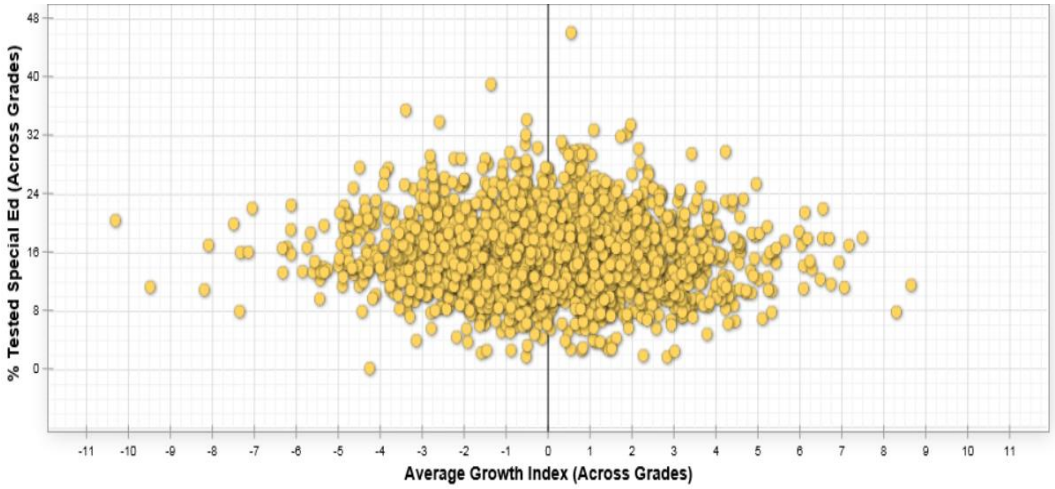
Pennsylvania has years of evidence to show that there are schools in Pennsylvania with a higher proportion of students with IEPS making growth. The same evidence can be found with schools with a high proportion of economically disadvantaged students and/or English Language Learners. There are also many schools with high achieving student yielding high growth. While achievement results may be highly correlated to demographics, growth results are typically not!

To see evidence of the scenarios with the percent of students with IEPs, look at the following example from 2014.

MATH



READING



What is most important to understand is that it is critical to have both types of data! Achievement data – tells you where the students are at a point in time AND Growth data – tells you how far the students have grown. Achievement and Growth – together they provide the more complete picture of student performance.

Misconception #8: PVAAS should always indicate growth if our percent of students proficient/advanced increased since last year.

PVAAS does NOT measure growth by students increasing or decreasing entire performance levels! PVAAS is sensitive to subtle changes in progress, even within performance levels. PVAAS value-added measures are NOT simple comparisons of test scores. With PVAAS, the entire history of students' test scores is used, not simply the two most recent scores.

Measuring Percent Proficient/Advanced by Grade Level

- When following the same grade level from one year to the next in determining the percent of students Proficient/Advanced, these are two different groups of students (i.e., 6th graders in 2013 are not the same group of students as 6th graders in 2014).
- As expected, different groups of students are different from year to year in terms of their average achievement!

Measuring Percent Proficient/Advanced by Cohort

- PVAAS looks at the most recent group of students and evaluates their growth from the prior school year in the prior grade level. Specifically, PVAAS is looking at the academic growth of that group of students in the most recent year they were tested on the PSSA and/or Keystone exams.
- This provides a view of the academic growth of the same students over time and across grade levels!

It is possible for PVAAS to indicate that students did not make a year's worth of academic growth even if the percent of Proficient and Advanced students is increasing (or even staying relatively stable) in a specified grade level or entire school.

- This may be because the percent of Proficient and Advanced students being compared is actually a comparison of two different groups of students.
- This may also be due to performance cut scores being at different percentiles of the state distribution.
- It is important to remember that PVAAS is NOT measuring growth by student groups increasing entire performance levels. Rather, PVAAS is more sensitive to growth even within those performance levels.
- You may find that the school has been successful in helping more students move from a non-proficient to a Proficient/Advanced level. However, at the same time, students already at the Proficient/Advanced level may be "slipping" in terms of their level of achievement compared to where they were the year prior (but still maintaining a Proficient/Advanced level overall).
- In other words, students may still be Proficient/Advanced, just not as high within those performance levels as they were in the prior year! Another, probably more common, scenario – students grow but do not grow enough to get to the next performance level. PVAAS will show the schools with these scenarios that their efforts are working even if more students have not yet crossed the bar of proficiency.

Example: Comparison of the SAME students from year to year.

		Grade 7 (2013-14)				
		Below Basic	Basic	Proficient	Advanced	
Grade 6 (2012-13)	Below Basic	6	35	8	1	71
	Basic	2	9	7	3	
	Proficient	1	4	41	8	79
	Advanced	0	5	10	10	
		69		88		

- In Grade 6 (2012-2013), 79 of the 150 (53%) students were Proficient or Advanced.
- In Grade 7 (2013-2014), 88 of the 150 (59%) students were Proficient or Advanced.

Misconception #9: PVAAS cannot measure the academic growth of districts and schools with high mobility rates.

Value-added analyses provide reliable and valid estimates of the effectiveness of districts and schools in supporting students to make one year's worth of academic growth, including those with high mobility!

- The use of the PAsecureID allows students' data to be matched longitudinally throughout students' K-12 academic careers, as long as the student has been enrolled and tested in Pennsylvania public schools. This PAsecureID stays with the student as they move from district to district, or school to school, across the Commonwealth.
- PVAAS analyses include ALL students, for which there are sufficient test data, including highly-mobile students.
- Policy decisions have been made to include ALL students in PVAAS district/school reporting who have been enrolled for a full academic year (on or before October 1) in PVAAS value-added analyses.
- From a statistical perspective, it is important to include highly-mobile students in the analysis because their exclusion could bias the results since this could be linked to lower achievement.
- From a philosophical perspective, as many students as possible must be included in the school's analysis to ensure that highly-mobile students receive the same level of attention as non-mobile students.

The EVAAS modeling approaches do take into account the quantity and quality of information available for each student! Student mobility is taken into account when students move from one school to the next from year to year. This mobility is taken into consideration when estimating value-added measures for each school in a particular year. Intra year student mobility is represented in the model by computing the mean of the population of students served. Students moving into and moving out of a school typically represent the same population of students.

Misconception #10: PVAAS cannot measure growth for groups of students who have missing data.

PVAAS analyses include ALL students, including those with some missing test data. This is one of the most inherent benefits of the PVAAS analyses.

- From a statistical perspective, it is important to include all students in the analyses because their exclusion could bias the resulting growth estimates.
- From a philosophical perspective, all students must be included in the district and school analyses to ensure that all students receive the same level of attention. This includes those students who may have missing data due to test records being lost, students moving from out of state, or students being sick on the day(s) the test is administered.
- Some subgroups of students tend to have more frequent missing data in their testing history, such as ELL students, migrant students, and students who are economically disadvantaged. The PVAAS model ensures that these students are included in the analyses, where other value-added models (such as a simple comparison of two test scores) would exclude these students.

For example, the PVAAS analyses for Math and English/Language Arts are based on the EVAAS multivariate response model (MRM) and use the correlation between current and previous scores in the non-missing data to estimate a mean for the previous and current score as if there were no missing data. This means that no values are explicitly imputed (statistically "made up") for the missing scores!

- This also means that no students are excluded from the analyses due to missing data!
- This method of addressing missing data has been shown to outperform other more simple methods of addressing missing data.
    - From Wright, S. P. (2004). "Advantages of a Multivariate Longitudinal Approach to Educational Value-Added Assessment Without Imputation." Paper presented at National Evaluation Institute, on-line at http://www.createconference.org/documents/archive/2004/Wright-NEI04.pdf.

Note: The approach to using all available data is similar for PVAAS for Science and Keystones.

Example illustrating the effect of missing data

Assume that ten students are given a test in two different years. The goal is to measure academic growth (gain) from one year to the next. The right side of the table shows what happens when some of the scores are missing. Two simple approaches to take when data are missing are to calculate the mean of the differences or to calculate the differences of the means.

Student	4^{th} Grade	5^{th} Grade	Gain
1	51.9	74.8	22.9
2	37.9	46.5	8.6
3	55.9	61.3	5.4
4	52.7	47.0	-5.7
5	53.6	50.4	-3.2
6	23.0	35.9	12.9
7	78.6	77.8	-0.8
8	61.2	64.7	3.5
9	47.3	40.6	-6.7
10	37.8	58.9	21.1
Mean	49.99	55.79	5.80
	Diff.	5.80	

Student	4^{th} Grade	5^{th} Grade	Gain
1	51.9		
2	37.9		
3	55.9	61.3	5.4
4	52.7	47.0	-5.7
5	53.6	50.4	-3.2
6	23.0	35.9	12.9
7		77.8	
8		64.7	
9	47.3	40.6	-6.7
10	37.8	58.9	21.1
Mean	45.01	54.58	3.97
	Diff.	9.57	

When there are no missing data, these two simple methods provide the same answer (5.80, shown in the left table). However, when there are missing data, each method provides a different result (9.57 vs. 3.97, shown in the right table). Missing data is very common to student testing data and must be taken into consideration, and thus a more sophisticated model is needed to address this problem. Some models throw out students with missing test scores; PVAAS does not. The sophisticated approach used by PVAAS estimates the means in each of these cells using relationships between students' test scores to obtain these means as if everyone was tested. In this way, the model finds unbiased growth measures without imputing any data. In fact, in practice, there is much more data to use to obtain these relationships.